

Qy 541 CTGAGCAATCCGAAGTACTCCAGCCTGCGGATCGTACGATCATTGAGCCGGACTCGCTG 600
Db 541 CTGAGCAATCCGAAGTACTCCAGCCTGCGGATCGTACGATCATTGAGCCGGACTCGCTG 600

Qy 601 CCAAACGCGGTACCAATATGAGCATTCAAGCGTGTGCAACGGCGGTGCCGTATTACGAG 660
Db 601 CCAAACGCGGTACCAATATGAGCATTCAAGCGTGTGCAACGGCGGTGCCGTATTACGAG 660

Qy 661 CAAGGCATCGAGTACCGCCTACGAAATTGCACGCCATTCCGAACGTGTACATCTACATG 720
Db 661 CAAGGCATCGAGTACCGCCTACGAAATTGCACGCCATTCCGAACGTGTACATCTACATG 720

Qy 721 GACGCCGCCACTCCGGCTGGCTTGGGTGCCAATAATGCCAGCGGATACGTACAGGAA 780
Db 721 GACGCCGCCACTCCGGCTGGCTTGGGTGCCAATAATGCCAGCGGATACGTACAGGAA 780

Qy 781 GTCCAGAAGGTCTCAACCGCGAGCATCGGGTCAACGGCATCGACGGCTCGTCACCAAC 840
Db 781 GTCCAGAAGGTCTCAACCGCGAGCATCGGGTCAACGGCATCGACGGCTCGTCACCAAC 840

Qy 841 ACGGCGAATTACACGCCGTTGAAGGAGCCGTTATGACCGCCACCCAGCAGGTCGGCGA 900
Db 841 ACGGCGAATTACACGCCGTTGAAGGAGCCGTTATGACCGCCACCCAGCAGGTCGGCGA 900

Qy 901 CAGCCGGTGGAGTCGGCAATTCTACCAAGTGGATCCTGACATCGACGAAGCCGACTAC 960
Db 901 CAGCCGGTGGAGTCGGCAATTCTACCAAGTGGATCCTGACATCGACGAAGCCGACTAC 960

Qy 961 GCGGTTGACTTGTACTCGCGCTCGTCGCCGCTGGCTTCCAAGCAGCAGCATGGCATGCTC 1020
Db 961 GCGGTTGACTTGTACTCGCGCTCGTCGCCGCTGGCTTCCAAGCAGCAGCATGGCATGCTC 1020

Qy 1021 ATCGACACCTTACGCAACGGTTGGGTGGTCCGAACGAACCAACAGGCCGAGCACCGCG 1080
Db 1021 ATCGACACCTTACGCAACGGTTGGGTGGTCCGAACGAACCAACAGGCCGAGCACCGCG 1080

Qy 1081 ACCGATGTCAACACCTCGTCAACCAGTCGAAGATTGACCTCGCAGCACCGCGCTG 1140
Db 1081 ACCGATGTCAACACCTCGTCAACCAGTCGAAGATTGACCTCGCAGCACCGCGCTG 1140

Qy 1141 TGGTGCAACCAGAACGGTGCAGGCCCTCGCCAGCCAGCGCAGGCAAGCCGACGGACTTC 1200
Db 1141 TGGTGCAACCAGAACGGTGCAGGCCCTCGCCAGCCAGCGCAGGCAAGCCGACGGACTTC 1200

Qy 1201 CCGAACCGCACCTCGACCGTATGTCTGGATCAAGCCCGGGTGAGTCGGACGGCACA 1260
Db 1201 CCGAACCGCACCTCGACCGTATGTCTGGATCAAGCCCGGGTGAGTCGGACGGCACA 1260

Qy 1261 AGCGCTGCGAGCGATCCGACAACGGCAAGAAGTCGGACCCATGTGCGACCCGACGTAC 1320
Db 1261 AGCGCTGCGAGCGATCCGACAACGGCAAGAAGTCGGACCCATGTGCGACCCGACGTAC 1320

Qy 1321 ACGACGTCGTACGGGTACTGACCAACCGTTACCGAACCTCCCGATGCCGCCAGTGG 1380
Db 1321 ACGACGTCGTACGGGTACTGACCAACCGTTACCGAACCTCCCGATGCCGCCAGTGG 1380

Qy 1381 TTCCCGGCCAGTTGACCAGCTGTCGCAACGCACGCCAGCGGTGCCGACGTCGACC 1440
Db 1381 TTCCCGGCCAGTTGACCAGCTGTCGCAACGCACGCCAGCGGTGCCGACGTCGACC 1440

Qy 1441 AGCTCGAGCCCGCCGCTCCGCCGAGTCGGTCCGCTCGCCGAGTCGGACCCGAGT 1500
Db 1441 AGCTCGAGCCCGCCGCTCCGCCGAGTCGGTCCGCTCGCCGAGTCGGACCCGAGT 1500

Qy 1501 CCGAGCCCGAGCAGCTGCCATGCCGTGCCGTCTCCGAGCTCGAGCCGTCCGTG 1560
Db 1501 CCGAGCCCGAGCAGCTGCCATGCCGTGCCGTCTCCGAGCTCGAGCCGTCCGTG 1560

Qy 1561 CCGAGCCCGAGTCCGAGCCGAGTAGCTGCCGTGCCGTCTCGAGCTCGAGCCGTCT 1620
Db 1561 CCGAGCCCGAGTCCGAGCCGAGTAGCTGCCGTGCCGTCTCGAGCTCGAGCCGTCT 1620

Qy 1621 CCGTCGCCGAGCCGAGTCCGAGCCGAGTAGCTGCCGTGCCGTCTCGAGCTCGAGC 1680
Db 1621 CCGTCGCCGAGCCGAGTCCGAGCCGAGTAGCTGCCGTGCCGTCTCGAGCTCGAGC 1680

Qy 1681 CCGTCTCCGTGCCGAGCCGAGTCCGAGCCGAGTAGCTGCCGTGCCGTCTCGACG 1740
Db 1681 CCGTCTCCGTGCCGAGCCGAGTCCGAGCCGAGTAGCTGCCGTGCCGTCTCGACG 1740

Qy 1741 TCGTCGCCGGTGTGGTGGCTGAAGGTGCAGTACAAGAACAAATGATTGGCGCCGGT 1800
Db 1741 TCGTCGCCGGTGTGGTGGCTGAAGGTGCAGTACAAGAACAAATGATTGGCGCCGGT 1800

Qy 1801 GATAACCAGATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGGTGGATTG 1860
Db 1801 GATAACCAGATCAAACCGGGTCTCCAGTTGGTGAATACCGGGTCGTCGTCGGTGGATTG 1860

Qy 1861 TCGACGGTGACGGTGCCTACTGGTCACCCGGATGGTGGTCGACACTGGTGTAC 1920
Db 1861 TCGACGGTGACGGTGCCTACTGGTCACCCGGATGGTGGTCGACACTGGTGTAC 1920

Qy 1921 AACTGTGACTGGCGCGATGGGTGTGGAAATATCCGCGCTCGTCGGCTCGGTGAAC 1980
Db 1921 AACTGTGACTGGCGCGATGGGTGTGGAAATATCCGCGCTCGTCGGCTCGGTGAAC 1980

Qy 1981 CCGGCGACGCCGACGGCGACACCTACCTGCAGTTGTCGTTACTGGTGGAACGTTGGCC 2040
Db 1981 CCGGCGACGCCGACGGCGACACCTACCTGCAGTTGTCGTTACTGGTGGAACGTTGGCC 2040

Qy 2041 GCTGGTGGTCGACGGGTGAGATTCAAACCGGGTGAATAAGAGTGACTGGTCGAATTTC 2100
Db 2041 GCTGGTGGTCGACGGGTGAGATTCAAACCGGGTGAATAAGAGTGACTGGTCGAATTTC 2100

Qy 2101 ACCGAGACCAATGACTACTCGTATGGGACGAACACCACTCCAGGACTGGACGAAGGTG 2160
Db 2101 ACCGAGACCAATGACTACTCGTATGGGACGAACACCACTCCAGGACTGGACGAAGGTG 2160

Qy 2161 ACGGTGTACGTCAACGGCGTGTGGTGTGGGGACTGAACCGTCCGGCACAGCCCCAGC 2220
Db 2161 ACGGTGTACGTCAACGGCGTGTGGTGTGGGGACTGAACCGTCCGGCACAGCCCCAGC 2220

Qy 2221 CCCACACCATCCCCGAGCCCGAGCCCGAGCCCGAGCCGGTGGGGATGTGACGCCGCC 2280
Db 2221 CCCACACCATCCCCGAGCCCGAGCCCGAGCCCGAGCCGGTGGGGATGTGACGCCGCC 2280

Qy 2281 AGTGTGCCGACCGGCTGGTGGTACGGGGTGAGTGGTCGTCGGTGTGCGTTGGCGTGG 2340
Db 2281 AGTGTGCCGACCGGCTGGTGGTACGGGGTGAGTGGTCGTCGGTGTGCGTTGGCGTGG 2340

Qy 2341 AATGCGTCACGGATAACGTGGGGTGCGCATTACAACGTGTACCGCAACGGGGTGTG 2400
Db 2341 AATGCGTCACGGATAACGTGGGGTGCGCATTACAACGTGTACCGCAACGGGGTGTG 2400

Qy 2401 GTGGGCCAGCCGACGGTGACCTCGTTACCGACACGGGTTGGCCCGGGAACCGCGTAC 2460
Db 2401 GTGGGCCAGCCGACGGTGACCTCGTTACCGACACGGGTTGGCCCGGGAACCGCGTAC 2460

Qy 2461 ACCTACACGGTGGCCCGGTGGACGCTGCAGTAACACCTCCGCCCATCCACCCCCGTC 2520
Db 2461 ACCTACACGGTGGCCCGGTGGACGCTGCAGTAACACCTCCGCCCATCCACCCCCGTC 2520

Qy 2521 ACCGCCACCAACGAGTCCCAGCCCCAGCCCCACGCCGACGGGACCACGGTCACCGAC 2580
Db 2521 ACCGCCACCAACGAGTCCCAGCCCCAGCCCCACGCCGACGGGACCACGGTCACCGAC 2580

Db 2521 ACCGCCACCACGAGTCCCAGCCCCAGCCCCACGGCAGGGGACCACGGTCACCGAC 2580
Qy 2581 TGCACGCCCGGTCTAACAGAATGGTGTGACAGCGTGCAGGGCAGAATACCGGTG 2640
Db 2581 TGCACGCCCGGTCTAACAGAATGGTGTGACAGCGTGCAGGGCAGAATACCGGTG 2640
Qy 2641 CAGACCAATGAGTGGAAATTCTCGTCGGCCAGCAGTGCCTCACCATCAATACCGCGACCGT 2700
Db 2641 CAGACCAATGAGTGGAAATTCTCGTCGGCCAGCAGTGCCTCACCATCAATACCGCGACCGT 2700
Qy 2701 GCCTGGACGGTGAGCACTGCGAACTTCAGCGGTGGGACCGGGCGTGCGCCCGACGTAT 2760
Db 2701 GCCTGGACGGTGAGCACTGCGAACTTCAGCGGTGGGACCGGGCGTGCGCCCGACGTAT 2760
Qy 2761 CCGTCGATCTACAAGGGCTGCCACTGGGCAACTGCACACAGAAGAACGTGGGATGCCG 2820
Db 2761 CCGTCGATCTACAAGGGCTGCCACTGGGCAACTGCACACAGAAGAACGTGGGATGCCG 2820
Qy 2821 ATTCAAGATCAGTCAGATTGGTCGGCTGTGACGTGTTGGAGTACGACGCAGGTGTCG 2880
Db 2821 ATTCAAGATCAGTCAGATTGGTCGGCTGTGACGTGTTGGAGTACGACGCAGGTGTCG 2880
Qy 2881 GGCGCGTATGACGTGGCCTACGACATTGGACGAACAGTACCCAACGACAACCGGTCA 2940
Db 2881 GGCGCGTATGACGTGGCCTACGACATTGGACGAACAGTACCCAACGACAACCGGTCA 2940
Qy 2941 CCAAACGGTACCGAAATCATGATTGGCTGAATTCGCTGGTGGGGTGCAGCCGTTGGG 3000
Db 2941 CCAAACGGTACCGAAATCATGATTGGCTGAATTCGCTGGTGGGGTGCAGCCGTTGGG 3000
Qy 3001 TCGCAGACAGCGACGGGTGTGACGGTCGCTGGTCACACGTGGAATGTCTGGCAGGGTCAG 3060
Db 3001 TCGCAGACAGCGACGGGTGTGACGGTCGCTGGTCACACGTGGAATGTCTGGCAGGGTCAG 3060
Qy 3061 CAGACCTCGTGGAAAGATTATTCTACGTCTGACCCCCGGTGCACGTCGATCAGTAAT 3120
Db 3061 CAGACCTCGTGGAAAGATTATTCTACGTCTGACCCCCGGTGCACGTCGATCAGTAAT 3120
Qy 3121 CTGGATTGAAGGCAGTTTCGCGGACGCCGCGGACCGCGGGTCGCTCAACACCTCCGAT 3180
Db 3121 CTGGATTGAAGGCAGTTTCGCGGACGCCGCGGACCGCGGGTCGCTCAACACCTCCGAT 3180
Qy 3181 TACCTGCTCGACGTTGAGGCCGGTTGAGATCTGGCAAGGTGGTCAGGGCTGGCAGC 3240
Db 3181 TACCTGCTCGACGTTGAGGCCGGTTGAGATCTGGCAAGGTGGTCAGGGCTGGCAGC 3240
Qy 3241 AACTCGTTCAGCGTCTCCGTGACGAGCGGACCGTCCAGCCGACACCGAGCCGAGCCCG 3300
Db 3241 AACTCGTTCAGCGTCTCCGTGACGAGCGGACCGTCCAGCCGACACCGAGCCGAGCCCG 3300
Qy 3301 ACGCCGACACCGAGCCGACGCCGACACCGTCTCCGAGCCGACCCCGTCGCCGAGTCG 3360
Db 3301 ACGCCGACACCGAGCCGACGCCGACACCGTCTCCGAGCCGACCCCGTCGCCGAGTCG 3360
Qy 3361 ACCAGCTGCCGTCTCGTGGGTGTGGCGTGCAGGGCAGTATGTGGTGAATAGTGAT 3420
Db 3361 ACCAGCTGCCGTCTCGTGGGTGTGGCGTGCAGGGCAGTATGTGGTGAATAGTGAT 3420
Qy 3421 TGGGGTTCTGGTTACGGCGACGGTGACGGTGACGAATACCGGGAGCCGGGACGAAC 3480
Db 3421 TGGGGTTCTGGTTACGGCGACGGTGACGGTGACGAATACCGGGAGCCGGGACGAAC 3480
Qy 3481 GGGTGGACGGTGGCGTGGTCGTTGGTGGGAATCAGACGGTCACGAACACTGGAACACT 3540
Db 3481 GGGTGGACGGTGGCGTGGTCGTTGGTGGGAATCAGACGGTCACGAACACTGGAACACT 3540
Qy 3541 GCGTTGACCCAATCAGGTGCATCGGTGACGGCGACGAACCTGAGTTACAACACGTGATC 3600

•
Db 3541 GCGTTGACCAATCAGGTGCATCGGTGACGGCGACGAACCTGAGTTACAACAACGTGATC 3600
Qy 3601 CAACCGGGTCAGTCGACCACCTCGGATTCAACGGAAGTTACTCAGGAACAAACGCCGCG 3660
Db 3601 CAACCGGGTCAGTCGACCACCTCGGATTCAACGGAAGTTACTCAGGAACAAACGCCGCG 3660
Qy 3661 CCGACGCTCAGCTGCACAGCCAGCTGA 3687
Db 3661 CCGACGCTCAGCTGCACAGCCAGCTGA 3687

Title: US-09-917-383-7

RESULT 2

US-09-917-384-7

; Sequence 7, Application US/09917384
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULASE FROM ACIDOTHERMUS
; TITLE OF INVENTION: CELLULOLYTICUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: GuxA

US-09-917-384-7

Query Match 100.0%; Score 1244; DB 23; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.3e-115;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
| | | | |
Qy 1 DCTPGPNQNGVTSVQDG**EYRVQT**NEWNSAQQCLTINTATGAWTVSTANFSGGTGGAPAT 60
| | | | |
Db 1 DCTPGPNQNGVTSVQDG**EYRVQT**NEWNSAQQCLTINTATGAWTVSTANFSGGTGGAPAT 60
| | | | |
Qy 61 YPSIYKGCHWGNC**TTK**NVG**MPIQ**ISQIGSAV**T**SWSTTQVSSGAYDVAYDIWTNSTPTTG 120
| | | | |
Db 61 YPSIYKGCHWGNC**TTK**NVG**MPIQ**ISQIGSAV**T**SWSTTQVSSGAYDVAYDIWTNSTPTTG 120
| | | | |
Qy 121 QPN**GTEIMIWLNSRGGVQ**PFGSQ**T**ATG**TVAGHT**WNW**QGQQT**SW**KIISYV**LTPGAT**SIS** 180
| | | | |
Db 121 QPN**GTEIMIWLNSRGGVQ**PFGSQ**T**ATG**TVAGHT**WNW**QGQQT**SW**KIISYV**LTPGAT**SIS** 180
| | | | |
Qy 181 NLDLKAI**FADAAARGSLNTSDY**LLDVEAG**FEI**W**QGGQGLGSNSFSV**SVTSG 231
| | | | |
Db 181 NLDLKAI**FADAAARGSLNTSDY**LLDVEAG**FEI**W**QGGQGLGSNSFSV**SVTSG 231

